Subscribe (Full Service)   Register (Limited Service, Free)   Login

Search: ⦿ The ACM Digital Library   ○ The Guide

+"reversed proxy" +"service time"

THE ACM DIGITAL LIBRARY

⁑ Feedback Report a problem Satisfaction survey

Terms used **reversed proxy service time**   Found **4** of **160,457**

Sort results by [relevance ▼]   🔖 Save results to a Binder   Try an Advanced Search
Display results [expanded form ▼]   ? Search Tips   Try this search in The ACM Guide
                                         ☐ Open results in a new window

Results 1 - 4 of 4

Relevance scale ☐ ▭ ▨ ▨ ▨

1 Aging through cascaded caches: performance issues in the distribution of web content   ▨
Edith Cohen, Haim Kaplan
August 2001 **ACM SIGCOMM Computer Communication Review , Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communications**, Volume 31 Issue 4
Full text available: 📄 pdf(327.13 KB)   Additional Information: full citation, abstract, references, citings, index terms

The Web is a distributed system, where data is stored and disseminated from both *origin* servers and *caches*. Origin servers provide the most up-to-date copy whereas caches store and serve copies that had been cached for a while. Origin servers do not maintain per-client state, and weak-consistency of cached copies is maintained by the origin server attaching to each copy an expiration time. Typically, the lifetime-duration of an object is fixed, and as a result, a copy fetched direc ...

2 The state of the art in locally distributed Web-server systems   ▭
Valeria Cardellini, Emiliano Casalicchio, Michele Colajanni, Philip S. Yu
June 2002 **ACM Computing Surveys (CSUR)**, Volume 34 Issue 2
Full text available: 📄 pdf(1.41 MB)   Additional Information: full citation, abstract, references, citings, index terms

The overall increase in traffic on the World Wide Web is augmenting user-perceived response times from popular Web sites, especially in conjunction with special events. System platforms that do not replicate information content cannot provide the needed scalability to handle large traffic volumes and to match rapid and dramatic changes in the number of clients. The need to improve the performance of Web-based services has produced a variety of novel content delivery architectures. This article w ...

Keywords: Client/server, World Wide Web, cluster-based architectures, dispatching algorithms, distributed systems, load balancing, routing mechanisms

3 A client-aware dispatching algorithm for web clusters providing multiple services   ▨
Emiliano Casalicchio, Michele Colajanni
April 2001 **Proceedings of the 10th international conference on World Wide Web**
Full text available: 📄 pdf(311.46 KB) Additional Information: full citation, references, citings, index terms

Keywords: clusters, dispatching algorithms, load balancing

4 Special issue on the PAPA 2002 workshop: Performance study of dispatching
algorithms in multi-tier web architectures

Mauro Andreolini, Michele Colajanni, Ruggero Morselli

September 2002 **ACM SIGMETRICS Performance Evaluation Review**, Volume 30 Issue 2

Full text available: pdf(1.16 MB)　　　　Additional Information: full citation, abstract, references, citings

The number and heterogeneity of requests to Web sites are increasing also because the Web
technology is becoming the preferred interface for information systems. Many systems
hosting current Web sites are complex architectures composed by multiple server layers with
strong scalability and reliability issues. In this paper we compare the performance of several
combinations of centralized and distributed dispatching algorithms working at the first and
second layer, and using different levels of st ...

Results 1 - 4 of 4

Useful downloads: Adobe Acrobat　QuickTime　Windows Media Player　Real Player

PORTAL
USPTO

Search: ⦿ The ACM Digital Library   ○ The Guide

+"service time" +monitor +proxy +uri

THE ACM DIGITAL LIBRARY

Feedback Report a problem Satisfaction survey

Terms used **service time** **monitor** **proxy** **uri**   Found **3** of **160,457**

Sort results by   [relevance ▼]
Display results   [expanded form ▼]

❖ Save results to a Binder
⍰ Search Tips
☐ Open results in a new window

Try an Advanced Search
Try this search in The ACM Guide

Results 1 - 3 of 3

Relevance scale ☐ ▦ ▦ ▦ ▦

1 The state of the art in locally distributed Web-server systems
Valeria Cardellini, Emiliano Casalicchio, Michele Colajanni, Philip S. Yu
June 2002 **ACM Computing Surveys (CSUR)**, Volume 34 Issue 2

Full text available: 📄 pdf(1.41 MB)   Additional Information: full citation, abstract, references, citings, index terms

The overall increase in traffic on the World Wide Web is augmenting user-perceived response times from popular Web sites, especially in conjunction with special events. System platforms that do not replicate information content cannot provide the needed scalability to handle large traffic volumes and to match rapid and dramatic changes in the number of clients. The need to improve the performance of Web-based services has produced a variety of novel content delivery architectures. This article w ...

Keywords: Client/server, World Wide Web, cluster-based architectures, dispatching algorithms, distributed systems, load balancing, routing mechanisms

2 Performance Workload Char. and Adaptation: Improving web performance by client characterization driven server adaptation
Balachander Krishnamurthy, Craig E. Wills
May 2002 **Proceedings of the 11th international conference on World Wide Web**

Full text available: 📄 pdf(241.76 KB)   Additional Information: full citation, abstract, references, citings, index terms

We categorize the set of clients communicating with a server on the Web based on information that can be determined by the server. The Web server uses the information to direct tailored actions. Users with poor connectivity may choose not to stay at a Web site if it takes a long time to receive a page, even if the Web server at the site is not the bottleneck. Retaining such clients may be of interest to a Web site. Better connected clients can receive enhanced representations of Web pages, such ...

Keywords: client characterization, client connectivity, server adaptation

3 Pushing politely: improving Web responsiveness one packet at a time
Brian D. Davison, Vincenzo Liberatore
September 2000 **ACM SIGMETRICS Performance Evaluation Review**, Volume 28 Issue 2

Full text available: 📄 pdf(650.86 KB)   Additional Information: full citation, abstract, citings, index terms

The rapid growth of traffic on the World-Wide Web results in heavier loads on networks and

servers and in increased latency experienced while retrieving web documents. This paper presents a framework that exploits idle periods to satisfy future HTTP requests speculatively and opportunistically. Our proposal differs from previous schemes in that speculative dissemination always gives precedence to on-demand traffic, uses ranged requests for improved performance, and can be implemented over a conn ...

Results 1 - 3 of 3

Subscribe (Full Service)   Register (Limited Service, Free)   Login

Search: ◉ The ACM Digital Library   ○ The Guide

+"service time" +monitor +proxy +url

THE ACM DIGITAL LIBRARY

Feedback Report a problem Satisfaction survey

Terms used **service time** **monitor** **proxy** **url**   Found **17** of **160,457**

Sort results by   [relevance ▼]   ◆ Save results to a Binder   Try an Advanced Search
                                          ? Search Tips   Try this search in The ACM Guide
Display results   [expanded form ▼]   ☐ Open results in a new window

Results 1 - 17 of 17

Relevance scale ☐ ▭ ▦ ▩ ▣

1 Performance issues of enterprise level web proxies
Carlos Maltzahn, Kathy J. Richardson, Dirk Grunwald
June 1997 **ACM SIGMETRICS Performance Evaluation Review , Proceedings of the
1997 ACM SIGMETRICS international conference on Measurement and
modeling of computer systems**, Volume 25 Issue 1

Full text available: 📄 pdf(1.75 MB)   Additional Information: full citation, abstract, references, citings, index terms

Enterprise level web proxies relay world-wide web traffic between private networks and the
Internet. They improve security, save network bandwidth, and reduce network latency.
While the performance of web proxies has been analyzed based on synthetic workloads,
little is known about their performance on real workloads. In this paper we present a study
of two web proxies (CERN and Squid) executing real workloads on Digital's Palo Alto
Gateway. We demonstrate that the simple CERN proxy architecture ...

2 The state of the art in locally distributed Web-server systems
Valeria Cardellini, Emiliano Casalicchio, Michele Colajanni, Philip S. Yu
June 2002 **ACM Computing Surveys (CSUR)**, Volume 34 Issue 2

Full text available: 📄 pdf(1.41 MB)   Additional Information: full citation, abstract, references, citings, index terms

The overall increase in traffic on the World Wide Web is augmenting user-perceived
response times from popular Web sites, especially in conjunction with special events.
System platforms that do not replicate information content cannot provide the needed
scalability to handle large traffic volumes and to match rapid and dramatic changes in the
number of clients. The need to improve the performance of Web-based services has
produced a variety of novel content delivery architectures. This article w ...

Keywords: Client/server, World Wide Web, cluster-based architectures, dispatching
algorithms, distributed systems, load balancing, routing mechanisms

3 Fast detection of communication patterns in distributed executions
Thomas Kunz, Michiel F. H. Seuren
November 1997 **Proceedings of the 1997 conference of the Centre for Advanced Studies
on Collaborative research**

Full text available: 📄 pdf(4.21 MB)   Additional Information: full citation, abstract, references, index terms

Understanding distributed applications is a tedious and difficult task. Visualizations based on

process-time diagrams are often used to obtain a better understanding of the execution of the application. The visualization tool we use is Poet, an event tracer developed at the University of Waterloo. However, these diagrams are often very complex and do not provide the user with the desired overview of the application. In our experience, such tools display repeated occurrences of non-trivial commun ...

4 Replication for web hosting systems

Swaminathan Sivasubramanian, Michal Szymaniak, Guillaume Pierre, Maarten van Steen
September 2004 **ACM Computing Surveys (CSUR)**, Volume 36 Issue 3

Full text available: pdf(374.99 KB)   Additional Information: full citation, abstract, references, index terms

Replication is a well-known technique to improve the accessibility of Web sites. It generally offers reduced client latencies and increases a site's availability. However, applying replication techniques is not trivial, and various Content Delivery Networks (CDNs) have been created to facilitate replication for digital content providers. The success of these CDNs has triggered further research efforts into developing advanced <i>Web replica hosting systems</i>. These are systems that ...

Keywords: Web replication, content delivery networks

5 Server performance and scalability: A method for transparent admission control and request scheduling in e-commerce web sites

Sameh Elnikety, Erich Nahum, John Tracey, Willy Zwaenepoel
May 2004 **Proceedings of the 13th international conference on World Wide Web**

Full text available: pdf(179.28 KB)   Additional Information: full citation, abstract, references, index terms

This paper presents a method for admission control and request scheduling for multiply-tiered e-commerce Web sites, achieving both stable behavior during overload and improved response times. Our method externally observes execution costs of requests online, distinguishing different request types, and performs overload protection and preferential scheduling using relatively simple measurements and a straight forward control mechanism. Unlike previous proposals, which require extensive changes to ...

Keywords: admission control, dynamic web content, load control, request scheduling, web servers

6 Cluster-based scalable network services

Armando Fox, Steven D. Gribble, Yatin Chawathe, Eric A. Brewer, Paul Gauthier
October 1997 **ACM SIGOPS Operating Systems Review , Proceedings of the sixteenth ACM symposium on Operating systems principles**, Volume 31 Issue 5
Full text available: pdf(2.42 MB)   Additional Information: full citation, references, citings, index terms

7 Web proxy caching: the devil is in the details

Ramón Cáceres, Fred Douglis, Anja Feldmann, Gideon Glass, Michael Rabinovich
December 1998 **ACM SIGMETRICS Performance Evaluation Review**, Volume 26 Issue 3

Full text available: pdf(539.43 KB)   Additional Information: full citation, abstract, citings, index terms

Much work in the analysis of proxy caching has focused on high-level metrics such as hit rates, and has approximated actual reference patterns by ignoring exceptional cases such as connection aborts. Several of these low-level details have a strong impact on performance, particularly in heterogeneous bandwidth environments such as modem pools connected to faster networks. Trace-driven simulation of the modem pool of a large ISP

suggests that "cookies" dramatically affect the cachability of resou ...

8 World Wide Web: Predicting web actions from HTML content

Brian D. Davison

June 2002 **Proceedings of the thirteenth ACM conference on Hypertext and hypermedia**

Full text available: pdf(243.13 KB)   Additional Information: full citation, abstract, references, index terms

Most proposed Web prefetching techniques make predictions based on the historical references to requested objects. In contrast, this paper examines the accuracy of predicting a user's next action based on analysis of the content of the pages requested recently by the user. Predictions are made using the similarity of a model of the user's interest to the text in and around the hypertext anchors of recently requested Web pages. This approa22ch can make predictions of actions that have never been ...

Keywords: WWW, information retrieval, prediction, prefetching, similarity, textual, user modeling

9 Special issue on the PAPA 2002 workshop: Performance study of dispatching algorithms in multi-tier web architectures

Mauro Andreolini, Michele Colajanni, Ruggero Morselli

September 2002 **ACM SIGMETRICS Performance Evaluation Review**, Volume 30 Issue 2

Full text available: pdf(1.16 MB)   Additional Information: full citation, abstract, references, citings

The number and heterogeneity of requests to Web sites are increasing also because the Web technology is becoming the preferred interface for information systems. Many systems hosting current Web sites are complex architectures composed by multiple server layers with strong scalability and reliability issues. In this paper we compare the performance of several combinations of centralized and distributed dispatching algorithms working at the first and second layer, and using different levels of st ...

10 Pervasive computing: Modeling service-based multimedia content adaptation in pervasive computing

Girma Berhe, Lionel Brunie, Jean-Marc Pierson

April 2004 **Proceedings of the 1st conference on Computing frontiers**

Full text available: pdf(691.71 KB)   Additional Information: full citation, abstract, references, index terms

Pervasive computing applications allow users to access information from anywhere while traveling and using variety of devices. Heterogeneity and limitation of resources involved in this application demand adaptation of content according to the current context (device, user, network etc.). The dynamic nature of adaptation mechanisms together with emerging opportunities of Web Service technology provides new approach of adaptation which is service-based. While this approach would provide a valuabl ...

Keywords: content adaptation services, media transformation, multimedia content delivery, pervasive computing

11 Adaptive push-pull: disseminating dynamic web data

Pavan Deolasee, Amol Katkar, Ankur Panchbudhe, Krithi Ramamritham, Prashant Shenoy

April 2001 **Proceedings of the 10th international conference on World Wide Web**

Full text available: pdf(152.08 KB)   Additional Information: full citation, references, citings, index terms

Keywords: World Wide Web, data dissemination, dynamic data, pull, push, resiliency, scalability, temporal coherency

12 Engineering server-driven consistency for large scale dynamic Web services
Jian Yin, Lorenzo Alvisi, Mike Dahlin, Arun Iyengar
April 2001 **Proceedings of the 10th international conference on World Wide Web**

Full text available: pdf(291.44 KB)   Additional Information: <u>full citation</u>, <u>references</u>, <u>citings</u>, <u>index terms</u>

Keywords: Web cache consistency, dynamic content, performance, scalability, volume lease

13 A client-aware dispatching algorithm for web clusters providing multiple services
Emiliano Casalicchio, Michele Colajanni
April 2001 **Proceedings of the 10th international conference on World Wide Web**

Full text available: pdf(311.46 KB)   Additional Information: <u>full citation</u>, <u>references</u>, <u>citings</u>, <u>index terms</u>

Keywords: clusters, dispatching algorithms, load balancing

14 Scheduling optimization for resource-intensive Web requests on server clusters
Huican Zhu, Ben Smith, Tao Yang
June 1999 **Proceedings of the eleventh annual ACM symposium on Parallel algorithms and architectures**

Full text available: pdf(1.19 MB)   Additional Information: <u>full citation</u>, <u>references</u>, <u>citings</u>, <u>index terms</u>

15 Learning response time for WebSources using query feedback and application in query optimization
Jean-Robert Gruser, Louiqa Raschid, Vladimir Zadorozhny, Tao Zhan
March 2000 **The VLDB Journal — The International Journal on Very Large Data Bases**,
Volume 9 Issue 1

Full text available: pdf(625.36 KB)   Additional Information: <u>full citation</u>, <u>abstract</u>, <u>citings</u>, <u>index terms</u>

The rapid growth of the Internet and support for interoperability protocols has increased the number of Web accessible sources, WebSources. Current wrapper mediator architectures need to be extended with a wrapper cost model (WCM) for WebSources that can estimate the response time (delays) to access sources as well as other relevant statistics. In this paper, we present a Web prediction tool (WebPT), a tool that is based on learning using query feedback from WebSources. The WebPT uses dimensions ...

Keywords: Data-intensive applications on the Web, Query languages and systems for Web data

16 Locality-aware request distribution in cluster-based network servers
Vivek S. Pai, Mohit Aron, Gaurov Banga, Michael Svendsen, Peter Druschel, Willy Zwaenepoel, Erich Nahum
October 1998 **Proceedings of the eighth international conference on Architectural support for programming languages and operating systems**, Volume 33 , 32

Issue 11 , 5

Full text available: pdf(1.59 MB)   Additional Information: <u>full citation</u>, <u>abstract</u>, <u>references</u>, <u>citings</u>, <u>index terms</u>

We consider cluster-based network servers in which a front-end directs incoming requests to one of a number of back-ends. Specifically, we consider *content-based request distribution;* the front-end uses the content requested, in addition to information about the load on the back-end nodes, to choose which back-end will handle this request. Content-based request distribution can improve locality in the back-ends' main memory caches, increase secondary storage scalability by partitioning th ...

17 Size-based scheduling to improve web performance

Mor Harchol-Balter, Bianca Schroeder, Nikhil Bansal, Mukesh Agrawal
May 2003 **ACM Transactions on Computer Systems (TOCS)**, Volume 21 Issue 2

Full text available: pdf(486.07 KB)   Additional Information: <u>full citation</u>, <u>abstract</u>, <u>references</u>, <u>citings</u>, <u>index terms</u>

Is it possible to reduce the expected response time of *every* request at a web server, simply by changing the order in which we schedule the requests? That is the question we ask in this paper.This paper proposes a method for improving the performance of web servers servicing static HTTP requests. The idea is to give preference to requests for small files or requests with short remaining file size, in accordance with the SRPT (Shortest Remaining Processing Time) scheduling policy.The imple ...

Keywords: Conservation law, SJF, SRPT, networking, scheduling, system performance and design, web servers

Results 1 - 17 of 17

Useful downloads: Adobe Acrobat   QuickTime   Windows Media Player   Real Player